DIFFERENT TYPES OF LEARNING (6 Pages)

BROAD TYPE OF MACHINE LEARNING

Supervised learning → For every instance, we tell what is the output.

- x, y (pre-classified training examples)
- Given an observation x, what is label for y?

x → y , also known as labelled data.
(Input) (output)

Unsupervised learning – x (Only given x, no labelled data i.e, No y is present).

- Given a set of x's, cluster to summarize them.
- Given data point, we may cluster them or summarize them or find some patterns.

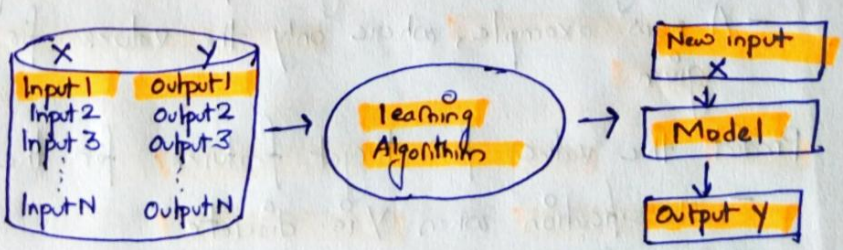
Reinforcement learning – Determine what to do based on rewards and punishments.

- Eg, Agent acting on an environment, and to figure out what action the agent must take at every step. The action that the agent take is based on the rewards or penalty is that the agents gets in different states.

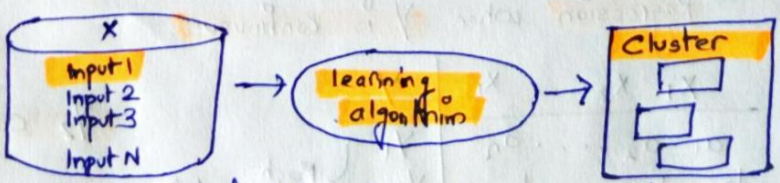
Semi-supervised learning – Combination of supervised and unsupervised learning.

- In data, we have some labelled data for training as well as unlabelled data.

Supervised Learning →



Unsupervised learning →



Based on similarity of data, it will group (cluster) data into different segment.

Semi Supervised learning →

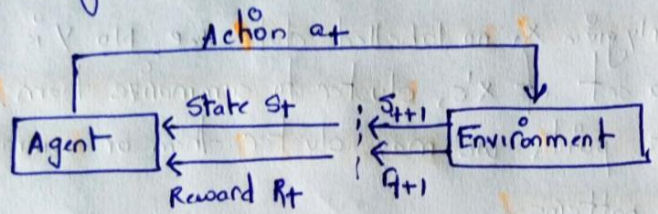
a) Labelled Data

b) Labelled & Unlabelled data

c) Supervised learning

d) Semi-supervised learning.

Reinforcement learning →



Agent is in the environment, it can take action in the environment which can impact the environment. In a particular stage, the agent takes an action and the environment goes to a new state and give some reward to the agent, that reward may be positive reward / negative reward / penalty / can be nothing at that particular time.

Supervised learning →

- A set of input features $X_1, \ldots, X_n$.
- A target feature Y.
- A set of training examples where the values for the input features and target features are given for each example.
- A new example, where only the values for the input features are given.

Predict the value for target features for the new example.

- Classification when Y is discrete.
- Regression when Y is continuous.

	X_1	X_2	...	X_n	Y	
I_1	a_1	a_2	...	a_n	Y_1	Each Instance (I_i) & traing instance
I_2	b_1	b_2	...	b_n	Y_2	
I_3	c_1	c_2	...	c_n	Y_3	
Test Instance	z_1	z_2	...	z_n	? Y_z.	

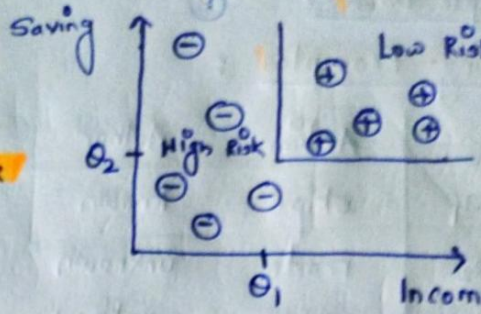
Y_z can be discrete or continuous

- Discrete (Classification) - Today it will Rain or not?
- Continuous (Regression) - Price of house

③

Examples of Classification –

- Credit Scoring
  - → 2 variables – Income and Saving
  - → To predict whether person is high risk person or low risk person.
  - → High risk person labelled with ⊖
    Low risk person labelled with ⊕



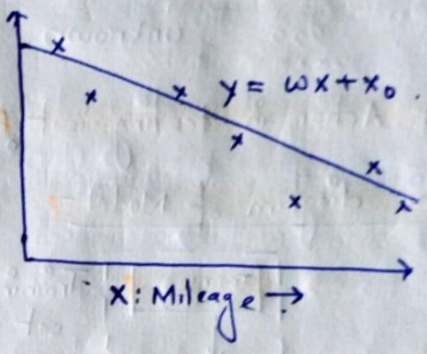
  - → Differentiate between low risk and high risk customers from their income and savings.
  - → Discriminant: if income $> \theta_1$ and saving $> \theta_2$ then low risk else high risk. (Rule / Classifier)

Example of Regression –

- Price of used car.
- 2 variables – x: car attribute
  y: price

generate a fn ← $y = g(x, \theta)$

$g()$ model,
θ parameters.



In regression, we come up with a function which take the input instances and parameters of the model.

Features –

- often, the individual observations are analyzed into set of quantifiable properties which are called features. May be.
  - ° Categorical → "A", "B", "AB" or "O" for blood type.
  - ° Ordinal → "large", "medium" or "small".
  - ° Integred value → Number of words in a text.
  - ° Real value → Height.

Example Data →

Training Examples :

features

	Action	Author	Thread	Length	Where
e_1	skips	known	new	long	Home
e_2	reads	unknown	new	short	Work
e_3	skips	unknown	old	long	Work
e_4	skips	known	old	long	home
e_5	reads	known	new	short	home
e_6	skips	known	old	long	work

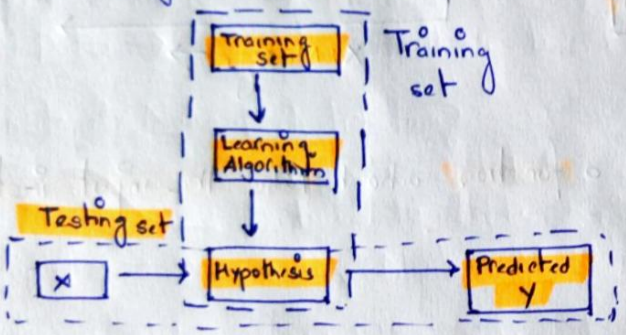
Values

New examples : Testing Set

e_7	???	known	new	short	work
e_8	???	unknown	new	short	work

Q) Predict Action in testing set ?

Schematic diagram of Model –



Training phase →

Testing phase →

⑥

Classification learning

- ① Task T:
  - Input: a set of instances $d_1, \ldots, d_n$.
    - An instance has a set of features.
    - We can represent an instances as a vector $d = \langle x_1, \ldots, x_n \rangle$
  - Output: a set of prediction $\hat{y}_1, \ldots, \hat{y}_n$
    - one of a fixed set of constant values:
      - $\{-1, +1\}$ or {cancer, healthy}, or $\{A, B, A^+, B^+\}$
- ② Performance metric P:
  - Probability (Right / wrong prediction) [We care about performance on the distribution, not the training data].
- ③ Experience E:
  - A set of labelled example (x, y) where y is true label for x.
  - Ideally, examples should be sampled from some fixed distribution D.

Hypothesis Space

- One way to think about a supervised learning machine is a device that explores a "hypothesis space".
  - Each setting of the parameters in the machine is a different hypothesis about the function that maps input vector to output vectors.
- And given the type of function, the hypothesis space is a set of candidate outputs that we get. So supervised learning is a set of functions which compromises the hypothesis space and we want to find out that function from hypothesis space which is most probable given you a training examples.

Terminology

- Features → The number of features or distinct traits that can be used to describe each item in a quantative manner.
- Feature vector → set of features, which make feature vector.
  → N dimensional vector of numerical feature that represent some object
- Instance space X → set of all possible object describable by features.
- Example (x, y) → Instance x with label $y = f(x)$.

Concept C - Subset of objects from X (C is unknown).
For examples, target is to find face of human being. So there will be many faces of objects like mole face, cars face, radio face. So, human face is a subset of Object X.

Target function F — Function which we are trying to learn. Whether it is a human face or not. It maps each instance $x \in X$ to target label $y \in Y$.

Examples (x,y) — Instances x with label $y = f(x)$.

Training Data S — Collection of examples observed by learning algorithm. Used to discover potentially predictive relationships.

— x — ^ — x —

INTERVIEW QUESTIONS.

Q) What is selection bias?
→ Selection of participant isn't random (method of collecting samples).
Type - i) Sampling bias → Non-random sample of population lead to less likely of the results.
ii) Time Interval → Terminated early / late.
iii) Data → Support hypothesis based on bad data.

Q) What is Normal distribution?
— 
Data is distributed around central value without any bias to left or right.
Properties — One mode, Mode = Median = Mean.

Q) Difference between Validation set and Test set?
— Validation set is considered as a part of training set as it is used for parameter selection and to avoid overfitting.
Test set is used for testing the performance of trained ML model.

Q) What is Interpolation and Extrapolation?
— Interpolation → Estimating a value from 2 known value from a list of values.
Extrapolation → It is approximating a value by extending a known set of values.